© Institut Pasteur/Elsevier
Paris 1996

Res. Microbiol.
1996, 147, 619-624

Genetic analysis of the replication region of the *Lactobacillus* plasmid vector pPSC22

P.S. Cocconcelli[1, 2] (*), M. Elli[2], B. Riboli[1] and L. Morelli[1, 2]

[1] *Centro Ricerche Biotecnologiche, Università Cattolica, via Milano 24, 26100, Cremona (Italy) and*
[2] *Istituto di Microbiologia, Università Cattolica, via Emilia Parmense 84, 29100 Piacenza (Italy)*

SUMMARY

The sequence and genetic organization of the 1,600-bp replication region of the *Lactobacillus* vector pPSC22, a plasmid derived from a 7-kb cryptic plasmid of *L. plantarum* used for the cloning of heterologous genes in several lactobacilli, were determined. Sequence analysis revealed the presence of a plus origin of replication containing the two functional elements *nic* and *bind*, required for initiation of the leading strands typical of the rolling circle (RC)-replicating plasmids belonging to the pLS1 family. Two open reading frames (*copA* and *repA*) were located within the *Lactobacillus* portion of pPSC22. The *repA* gene product, a 234-amino acid protein, showed homologies with the Rep protein of the streptococcal plasmid pLS1 and contained the three conserved domains detected in most Rep proteins of RC-replicating plasmids and ss-coliphages. The genetic organization of the replication region of pPSC22 shared relevant homologies with the lactococcal plasmids pWV01 and pFX2.

Key-words: Plasmid, *Lactobacillus*; Plasmid vectors, Origin of replication, RC, Nucleotide sequence.

INTRODUCTION

Lactobacilli have been used in food production for centuries and, because of their health properties, were introduced into drugs decades ago. This group of bacteria, because of their industrial significance, has become the focus of a rapidly increasing number of genetic studies especially devoted to the application of recombinant DNA techniques for the genetic modification of *Lactobacillus* strains. Plasmids (1.5 to 60 kb in size) are frequently found in various strains of *Lactobacillus* and recently, the mode of replication of some of them has been studied in detail. DNA sequence analysis of small multicopy replicons from *L. plantarum*, (Bates and Gilbert, 1989; Bouia *et al.*, 1989; Skaugen, 1989; Leer *et al.*, 1992; Vujcic and Topisirovic, 1993), *L. pentosus* (Leer *et al.*, 1992) and *L. hilgardii* (Josson *et al.*, 1989) has shown that they belong to a family of highly interrelated plasmids replicating via a similar mechanism, termed the rolling circle (RC) (Novick, 1989; Gruss and Ehrlich, 1989), already observed in a number of Gram-positive multicopy plasmids.

Submitted November 29, 1995, accepted April 9, 1996.

(*) Corresponding author.

A 1.6-kb fragment of a 7-kb cryptic plasmid of *L. plantarum*, containing the functions for autonomous replication, was used to construct the plasmid vector pPSC22 for lactobacilli (Cocconcelli *et al.*, 1991). This vector was shown to replicate and was used for the cloning of heterologous genes in a variety of *Lactobacillus* species, in *Lactococcus lactis* subsp. *lactis*, in *Bacillus subtilis*, in *Ruminococcus albus* and in Gram-negative bacteria, including *Escherichia coli* (Cocconcelli *et al.*, 1991, 1992). It was also demonstrated that pPSC22 replicates via a single-stranded DNA intermediate and shows homologies with the origin of replication of other Gram-positive RCR plasmids such as pE194 (Horinouchi and Weisblum, 1982). The genetic organization of several RC-replicating plasmids is now well established (Gruss and Ehrlich, 1989; del Solar *et al.*, 1993b; Espinosa *et al.*, 1995), and on the basis of homologies in the leading strand initiation and control region, four families have been established (Novick, 1989).

In this study, we describe the sequence and the genetic organization of the origin of replication of the *Lactobacillus* vector pPSC22.

MATERIALS AND METHODS

Plasmids and organisms, and growth conditions

The *E. coli* JM109 strain and the plasmids pUC18 and pUC19 (Yanisch-Perron *et al.*, 1985) were used for subcloning of the *Taq*I-fragments containing the origin of replication of pPSC22. The hosts of pPSC22 were the *E. coli* strain HB101 (Boyer and Roulland-Dussoix, 1969), and *L. reuteri* DSM20016. MRS medium (Difco) was used for culturing *L. reuteri*.

Determination of the nucleotide sequence

Plasmid pPSC22 was digested with the restriction endonuclease *Taq*I, and 520- and 1080-bp fragments were cloned into the *Acc*I site of plasmid pUC19 and transformed in *E. coli* JM109. Overlapping templates were obtained by exonuclease III digestion using the Erase-a-Base™ system (Promega). The complete nucleotide sequence of both strands was determined using the Taq dye primer cycle and Taq DyeDeoxy™ terminator cycle kits (Perkin Elmer) following the supplier's instructions. Sequence analysis was obtained with the Applied Biosystem 373A DNA sequencer. The software package of the University of Wisconsin Genetic Group was used for the DNA sequence analyses.

Plasmid transformation

Plasmid DNA was transformed in *E. coli* and in *Lactobacillus* and by means of electroporation as already reported (Dower *et al.*, 1988; Cocconcelli *et al.*, 1991).

General DNA manipulations

Plasmid DNA was extracted from lactobacilli as previously reported (Morelli *et al.*, 1987). Other methods of DNA manipulation were as reported by Sambrook *et al.* (1989).

RESULTS

Nucleotide sequence of the replication region of pPSC22

The nucleotide sequence of the 1,600-bp fragment containing the replication functions of pPSC22 is reported in figure 1. The calculated G+C% of the *Lactobacillus* portion of pPSC22 is 35.1%, lower than the 44-46% range reported for *L. plantarum* species (Kandler and Weiss, 1986). Strong DNA sequence homologies were found with the plasmids belonging to the family pLS1/pE194 of RC plasmids (del Solar *et al.*, 1993b). Comparison of the sequence data of the *Lactobacillus* portion of pPSC22 with the replication regions of the lactococcal plasmids pWV01 (Leenhouts *et al.*, 1991) and pFX2 (Xu *et al.*, 1991) demonstrated that the replication

ORF = open reading frame.
RBS = ribosome-binding site.
RC = rolling circle (replication mechanism).

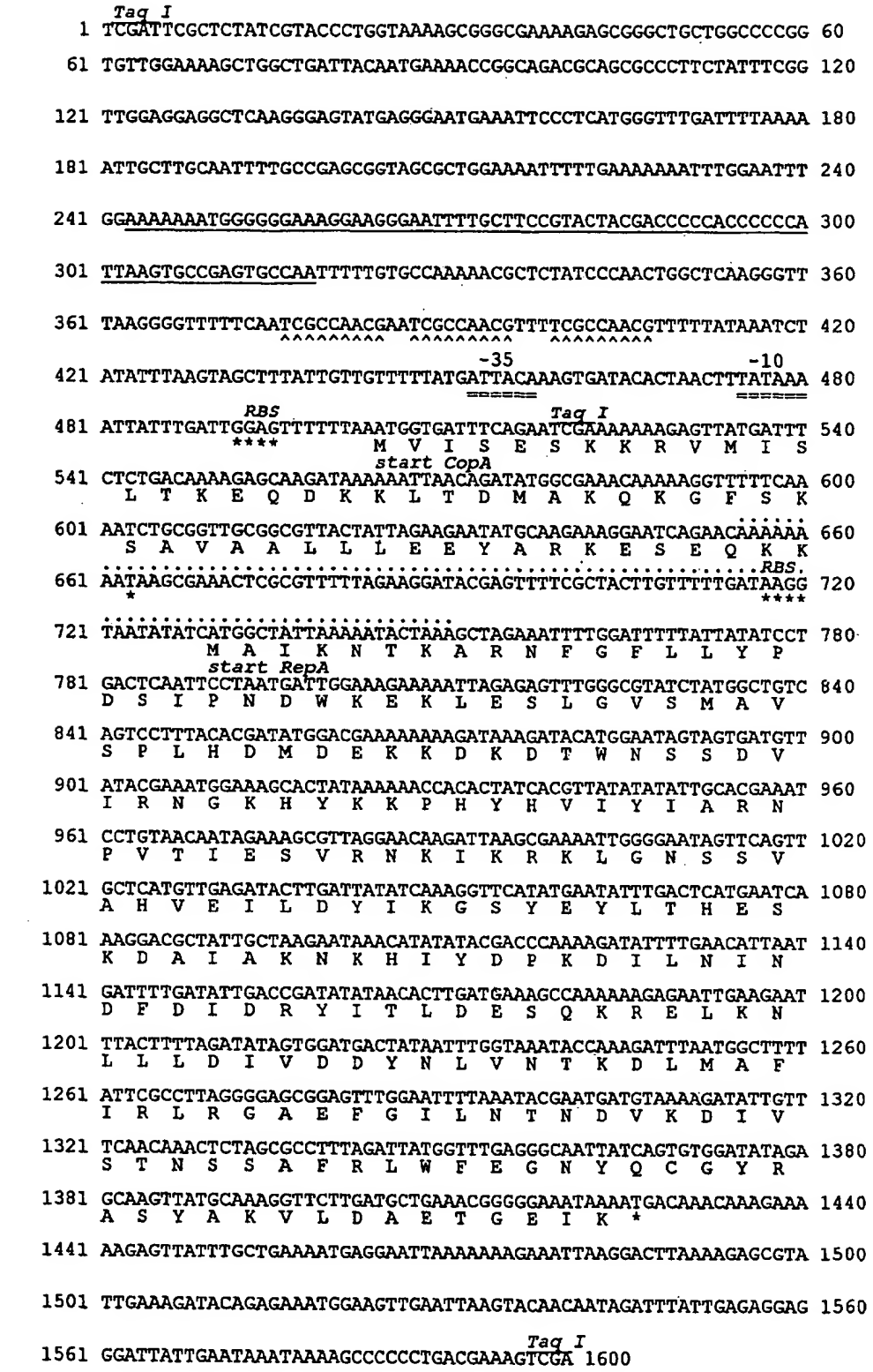

Fig. 1. Nucleotide sequence of the *Lactobacillus* portion (*Taq*I-*Taq*I-*Taq*I) of pPSC22.

The deduced amino acid sequences of the potential ORFs are shown. Double-underlined regions and asterisks indicate potential promoters and ribosome binding sites. A possible *dso* site (underlined) and iterons (inverted letter v) are indicated. The possible area coding for the CtRNA is marked by a dotted line. These sequence data were assigned accession no. X95843 in the EMBL data library.

functions of these plasmids are nearly identical (96% sequence homology).

pPSC22 double-stranded origin of replication

The essential region containing the elements required for *in cis* initiation of the leading strand synthesis (*dso*) was identified within the 1,600-bp fragment of pPSC22. The location of the *dso* in pPSC22, previously identified in the 520-kb *Taq*I fragment through complementation experiments (Cocconcelli *et al.*, 1991), was confirmed by deleting with exonuclease III and transforming into *L. reuteri* DSM20016 and *E. coli* HB101. Deletions resulted in the plasmid no longer being functional in lactobacilli and in the Gram-negative host. By sequence comparison, the two regions, *bind* and *nic*, composing the *dso* of the pLS1 family, were found in the 520-bp *Taq*I fragment of pPSC22. The *nic* site was identified as a region of dyad symmetry (positions 243-317) potentially capable of forming a stem-loop structure ($\Delta G^\circ = -22$ kcal/mol) containing, on the side of the stem, the possible site of DNA nicking TACTACG\A as described for pLS1 (de la Campa *et al.*, 1990). Downstream from the palindromic structures of the *nic* site, iterated sequences, potentially forming the *bind* locus involved in the binding of the initiation protein (de la Campa *et al.*, 1990; del Solar *et al.*, 1993a), were present. The three iterons (TCGCCAACG) found in the sequences of the *dso* of pPSC22 showed a high degree of homology with iterated sequences of the pLS1 family, as reported by del Solar *et al.* (1993a). In addition, the motif GTTT found in pLS1, pWVO1, pFX2 and pKMK1 was located in the direct repeats of the *bind* site of pPSC22.

Plasmid replication proteins

Downstream from the *dso*, computer-assisted analysis revealed the presence of two open reading frames, ORF1 (positions 504-662) and ORF2 (positions 730-1425), potentially capable of encoding polypeptides of 53 and 234 amino acids, respectively. ORF1 was preceded by a typical ribosome-binding site (RBS) (GGAG) 8 bp upstream from the translation initiation codon ATG. Ten bp upstream from the ATG of ORF2, another possible RBS (AAGG) was located. One putative promoter structure (ATTACA-17nucleotides-TATAAA) was present 23 bp upstream from the ATG of ORF1. The putative product of ORF1 (*copA* gene), a peptide containing 53 amino acids, showed a high degree of homology with the Cop transcription repressor proteins of the RC plasmids of the pLS1 family. The predicted secondary structures showed an α-helix-turn-α-helix composed of two helixes from residues 16-27 and between residues 34-52, and the turn motif present in the spacing region (28-33) between the helixes. The 26.8-kDa peptide (RepA), coded by ORF2, showed strong homologies with the replication initiation proteins (Rep) of the pLS1 family. Five conserved motifs, R-I to R-V, as described by del Solar *et al.* (1993b), and two conserved H residues were found in the amino acid sequence of RepA.

Possible regulative mechanism of pPSC22 replication

In the sequence of the replication region of pPSC22, the two elements involved in the control of replication and copy number present in the pLS1 family were found: the CopA protein, as described above, and a small antisense RNA that overlapped the translation initiation signal of the *repA* gene (positions 750-654). The area coding for the countertranscribed RNA (CtRNA) is shown in figure 1. A putative promoter from position 750 to position 721 (−35 region TTTAGT-17-bp spacing, −10 region TATATT) was found. Downstream from the promoter region, inverted repeats were present, forming a possible transcription terminator.

DISCUSSION

In the study described here, the nucleotide sequence of the 1600-bp replication region of the *Lactobacillus* plasmid vector pPSC22 was determined and analysed. On the basis of sequence

analysis of the leading strand initiation and control region, the *Lactobacillus* vector pPSC22 was demonstrated to belong to the recently described pLS1 family (del Solar *et al.*, 1993a) of highly interrelated plasmids.

The genetic elements involved in the initiation of replication and in the copy number control of the pLS1 family, the loci *bind* and *nic* of *dso* (del Solar *et al.*, 1993a), the *repA* gene, the *copA* gene, coding for the Cop repressor protein (del Solar and Espinosa 1992), and a CtRNA complementary to the *cop-rep* mRNA (Novick, 1989), were found in the sequence of pPSC22.

Comparison of the sequence data of the *Lactobacillus* portion of pPSC22 with the replication regions of the lactococcal plasmids pWVO1 (Leenhouts *et al.*, 1991) and pFX2 (Xu *et al.*, 1991), a vector derived from the plasmid pD125 (Xu *et al.*, 1990), demonstrated that the replication functions of these plasmids are nearly identical (96% sequence homology). The presence, in the highly interrelated family of plasmids that replicate via an RC mechanism, of common cassettes of genetic information was already reported (Projan and Novick, 1988). The high degree of homology among replication regions of plasmids of different molecular mass and isolated from different bacterial hosts (pPSC22 was derived from a 7-kb cryptic plasmid of *L. plantarum*, pWVO1 was a native 2.2-kb plasmid of *Lactococcus lactis* subsp. *cremoris*, and pFX2 was obtained from pD125, a 5.5-kb plasmid of *L. lactis* subsp. *lactis*) is an example of blocks of genetic information spread among microorganisms.

Acknowledgements

This work was supported by the National Research Council of Italy (CNR), Special Project RAISA, Subproject N. 4 Paper N.

Analyse génétique de la région de réplication du vecteur plasmidique pPSC22 de *Lactobacillus*

Nous avons déterminé la séquence et l'organisation génétique de la région de réplication, de 1.600 bp, du vecteur plasmidique pPSC22 de *Lactobacillus*, un plasmide dérivé d'un plasmide cryptique de 7 kb de *L. plantarum* utilisé pour le clonage des gènes hétérologues chez divers lactobacilles. L'analyse séquentielle révèle la présence d'une origine de réplication « plus » contenant les deux éléments fonctionnels, *nic* et *bind*, requis pour l'initiation observée avec les plasmides à réplication «rolling circle» qui appartiennent à la famille pLS1. Deux ORFs (*copa* et *repA*) ont été localisés dans la portion *Lactobacillus* de pPSC22. Le produit du gène *repA*, une protéine de 234 acides aminés, a des homologies avec la protéine Rep du plasmide streptococcique pLS1 et contient les trois domaines conservés que l'on décèle dans la plupart des protéines Rep des plasmides à réplication «rolling circle» et chez les coliphages ss. L'organisation de la région de la réplication de pPSC22 montre des homologies significatives avec les plasmides lactococciques pWVO1 et pFX2.

Mots-clés: Plasmide, *Lactobacillus*; Plasmides vecteurs, Origine de réplication, RC, Séquences nucléotidiques.

References

Bates, E.M. & Gilbert, H.J. (1989), Characterization of a cryptic plasmid from *Lactobacillus plantarum*. *Gene*, 85, 253-258.

Bouia, A., Bringel, F., Frey, L., Kammer, B., Belarbi, A., Guyonvarch, A. & hubert, J.C. (1989), Structural organization of pLP1, a cryptic plasmid from *Lactobacillus plantarum* CCM 1904. *Plasmid*, 22, 185-192.

Boyer, H.W. & Roulland-Dussoix, D. (1969), A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. *J. Mol. Biol.*, 41, 459-472.

Cocconcelli, P.S., Gasson, M.J., Morelli, L. & Bottazzi, V. (1991), Single-stranded DNA plasmid, vector construction and cloning of *Baccilus strearothermophilus* α-amylase in *Lactobacillus*. *Res. Microbiol.*, 142, 643-652.

Cocconcelli, P.S., Ferrari, E., Rossi, F. & Bottazzi, V. (1992), Plasmid transformation of *Ruminococcus albus* by means of high-voltage electroporation. *FEMS Microbiol. Lett.*, 94, 203-207.

De la Campa, A.G., Del Solar, G.H. & Espinosa, M. (1990), Initiation of replication of plasmid pLS1. The initiatior protein repB acts on two distant DNA regions. *J. Mol. Biol.*, 213, 247-262.

Del Solar, G. & Espinosa, M. (1992), The copy number of plasmid pLS1 is regulated by two trans-acting plasmid products: the antisense RNA and the repressor Rep A. *Mol. Microbiol.*, 6, 83-94.

Del Solar, G., Moscoso, M. & Espinosa, M. (1993a), *in vivo* definition of the functional origin of replication (ori(+)) of the promiscuous plasmid pLS1. *Mol. Gen. Genet.*, 237, 65-72.

Del Solar, G., Moscoso, M. & Espinosa, M. (1993b), Rolling circle-replicating plasmids from Gram-positive and Gram-negative bacteria: a wall falls. *Mol. Microbiol.*, 8, 789-796.

Dower, W.J., Miller, J.F. & Rangsdale, C.W. (1988), High efficiency transformation of *Escherichia coli* by high voltage electroporation. *Nucl. Acids. Res.*, 16, 6127-6145.

Espinosa, M., Del Solar, G., Rojo, F. & Alonso, J.C. (1995), Plasmid rolling circle replication and its control. *FEMS Microbiol. Lett.*, 130, 111-120.

Gruss, A. & Ehrlich, S.D. (1989), The family of highly interrelated single-stranded deoxyribonucleic acid plasmids. *Microbiol. Rev.*, 53, 231-241.

Horinouchi, S. & Wiesblum, B. (1982), Nucleotide sequence and functional map of pE194, a plasmid that specifies inducible resistance to macrolide, lincosamide, and streptogramin type B antibiotics. *J. Bacteriol.*, 150, 804-814.

Josson, K., Soetaert, P., Michielis, F., Joos, J. & Mahillon, J. (1990), *Lactobacillus hilgardii* plasmid pLAB1000 consists of two functional cassettes commonly found in other Gram-positive organisms. *J. Bacteriol.*, 172, 3089-3099.

Kandler, O. & Weiss, N. (1986), Regular, non-sporing Gram-positive rods, *in* "Bergey's manual of systematic bacteriology" (P.H. Sneath Ed.) (pp. 1208-1260). Williams & Wilkins, Baltimore, MD.

Leenhouts, K.J., Tolner, B., Bron, S., Kok, J., Venema, G. & Seegers, J.F.M. (1991), Nucleotide sequence and characterization of the broad-host-range lactococcal plasmid pWVO1. *Plasmid*, 26, 55-66.

Leer, R.J., Van Luijk, N., Posno, M. & Powels, P.H. (1992), Structural and functional analysis of two cryptic plasmids from *Lactobacillus pentosus* MD353 and *Lactobacillus plantarum* ATCC 8014. *Mol. Gen. Genet.*, 234, 265-274.

Morelli, L., Cocconcelli, P.S., Bottazzi, V., Damiani, G., Ferretti, L. & Sgaramella, V. (1987), *Lactobacillus* protoplast transformation. *Plasmid*, 17, 73-75.

Novick, R.P. (1989) Staphylococcal plasmids and their replication. *Annu. Rev. Microbiol.* 43, 537-565.

Projan, S.J. & Novick, R.P. (1988), Comparative analysis of five related staphylococcal plasmids. *Plasmid*, 19, 203-221.

Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989), Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Skaugen, M. (1989), The complete nucleotide sequence of a small cryptic plasmid from *Lactobacillus plantarum. Plasmid*, 22, 175-179.

Vujcic, M., Topisirovic, L. (1993), Molecular analysis of the rolling-circle replicating plasmid pA1 of *Lactobacillus plantarum* A112. *Appl. Environ. Microbiol.*, 59, 274-280.

Yanisch-Perron, C., Vieira, J. & Messing, J. (1985), Improved M13 phage cloning vectors and host strains: nucleotide sequence of the M13mp18 and pUC19 vectors. *Gene*, 33, 103-109.

Xu, F., Pearce, L.E. & Yu, P.L. (1990), Molecular cloning of a proteinase gene from *Lactococcus lactis* subsp. *cremonis* and construction of a new lactococcal vector pFX1. *Arch. Microbiol.*, 154, 99-104.

Xu, F., Pearce, L.E. & Yu, P.L. (1991), Genetic analysis of a lactococcal plasmid replicon. *Mol. Gen. Genet.*, 227, 33-39.